

EXHIBIT B

1 COOLEY LLP
BOBBY GHAJAR (198719)
2 (bghajar@cooley.com)
COLETTE GHAZARIAN (322235)
3 (cghazarian@cooley.com)
1333 2nd Street, Suite 400
4 Santa Monica, California 90401
Telephone: (310) 883-6400
5 Facsimile: (310) 883-6500

6 MARK WEINSTEIN (193043)
(mweinstein@cooley.com)
7 KATHLEEN HARTNETT (314267)
(khartnett@cooley.com)
8 JUDD LAUTER (290945)
(jlauter@cooley.com)
9 3175 Hanover Street
Palo Alto, CA 94304-1130
10 Telephone: (650) 843-5000
Facsimile: (650) 849-7400
11

Counsel for Defendant Meta Platforms, Inc.

12
13 **UNITED STATES DISTRICT COURT**
14 **NORTHERN DISTRICT OF CALIFORNIA**

15 RICHARD KADREY, et al.,
16 Individual and Representative Plaintiffs,
17 v.
18 META PLATFORMS, INC., a Delaware
19 corporation;
20 Defendant.

Case No. 3:23-cv-03417-VC

**DEFENDANT META PLATFORMS, INC.'S
RESPONSES AND OBJECTIONS TO
PLAINTIFFS' FIRST SET OF
INTERROGATORIES**

Trial Date: None
Date Action Filed: July 7, 2023

the Instruction if overly broad, as well. Subject to any objections applicable to a particular Interrogatory, Meta will conduct a reasonable, proportionate search for non-privileged, relevant, responsive information within its possession, custody, or control.

17. In responding to all Interrogatories, Meta will comply with the requirements of the Federal Rules of Evidence and Federal Rule of Civil Procedure 26.

III. OBJECTIONS AND RESPONSES TO INDIVIDUAL INTERROGATORIES

INTERROGATORY NO. 1:

Describe in detail the data You have used to train or otherwise develop the Meta Language Models, Including, for each:

- a. How You obtained the data, e.g., by scraping the data, purchasing it from third parties, or by other means;
- b. All sources of Data, including any third parties that provided data sets;
- c. To the extent the data was derived from publicly available websites, a list of all such websites and, for each, the percentage of the data corpus that is derived from that website;
- d. The categories of content included in the data and the extent to which each category is represented in the data corpus (i.e., as a percentage of data used to train the model);
- e. All policies and procedures Related to identifying, assessing, vetting and selecting sources of data for the model.

RESPONSE TO INTERROGATORY NO. 1:

Meta incorporates by reference its objections and definitions above, including to the terms “You” and “Meta Language Models.” Meta further notes that the capitalized term “Related” is not defined; Meta construes that term coextensively with “concerning.”

As an initial matter, Meta objects to this Interrogatory because it consists of multiple, separate Interrogatories, each which count toward Plaintiffs’ limit under the Federal Rules. For example, the question about what data used to train a model is separate from how it was obtained, and further, subparts (d) and (e) are not subsumed within and necessarily related to the primary question, and purport to require a calculation of percentages of data, and separate identification of “policies” and “procedures” for (1) identifying, (2) assessing, (3) vetting, and (4) selecting data.

1 This Interrogatory consists of *at least* three Interrogatories, and depending on how it is interpreted,
2 many more. In answering the Interrogatory, Meta does not waive this objection.

3 Meta objects to this Interrogatory because, on its face, it does not exclude legal advice or
4 opinions, which are subject to attorney-client privilege and/or attorney work product doctrine, in
5 particular as to subpart (e). Meta will not produce privileged materials or attorney work product.

6 Meta objects to this Interrogatory as vague and ambiguous as to the term “data,” which is
7 undefined. Meta will construe “data” to mean Training Data (as construed above).

8 Meta objects to this Interrogatory as vague, ambiguous, and unintelligible as to “percentage
9 of that data corpus that is derived from that website” because “data corpus” is undefined, and Meta
10 is accordingly unable to interpret and respond to subpart (c). Even if “data corpus” were defined,
11 the subject matter of subpart (c) would be overbroad, unduly burdensome, and disproportionate to
12 the needs of the case and seeks information that is not relevant to the parties’ claims and defenses.
13 Meta will not respond to subpart (c).

14 Meta objects to the undefined phrase “categories of content, which is vague, ambiguous,
15 and unintelligible.

16 Meta objects to this Interrogatory to the extent that it seeks information that is not within
17 Meta’s possession, custody, or control.

18 Subject to and without waiving the foregoing objections, and pursuant to the terms of the
19 Protective Order and the ESI Order, Meta responds as follows: Meta incorporates by reference the
20 identification of datasets used to train Llama 1 that is included in the publicly available paper
21 “LLaMA: Open and Efficient Foundation Language Models.” Such datasets were used to train
22 Llama 1. Meta will produce a copy of that paper in its forthcoming production pursuant to Rule
23 33(d).

24 Meta will conduct a reasonable search for additional non-privileged information or, in
25 accordance with Rule 33(d), documents in Meta’s possession, custody, or control, sufficient to
26 show any other datasets used to train the Meta Language Models (as construed above), as well as
27 policies and procedures for identifying, assessing, vetting, and selecting sources of data for those
28 models.

1 Discovery is ongoing and Meta will also supplement its response to this Interrogatory to
 2 identify the sources of such datasets and general categories of data within them, to the extent that
 3 such information is within Meta's possession, custody, or control.

4 Discovery is continuing and Meta reserves the right to supplement or amend its response at
 5 a later time.

6 **INTERROGATORY NO. 2:**

7 Describe in detail the process for modifying a Meta Language Model in order to create a
 8 variant of the model. Include in Your response:

9 a. The circumstances under which You opt to modify a Meta Language Model to create
 10 a variant;

11 b. The individual(s), type(s) of individual by title and area of expertise, or department
 12 responsible for the modifications;

13 c. The steps taken to modify the model;

14 d. How You determine whether the model has been adequately modified; and

15 e. A description of any modifications You actually undertook in order to correct or
 16 remediate any Meta Language Model's propensity to emit protected expression from its Training
 17 Data.

18 **RESPONSE TO INTERROGATORY NO. 2:**

19 Meta incorporates by reference its objections and definitions above, including to the terms
 20 "Meta Language Models," "You," and "Training Data."

21 As an initial matter, Meta objects to this Interrogatory because it consists of multiple,
 22 separate Interrogatories, each which count toward Plaintiffs' limit under the Federal Rules. For
 23 example, the question about the process of "modifying" a particular language model is separate
 24 from subpart (b), which asks a question about the specific expertise of each individual who might
 25 have been "responsible" for any modifications. Moreover, subpart e asks about steps taken to
 26 "correct or remediate" *any* language model's "propensity" to emit "protected expression," which is
 27 a separate inquiry altogether. This Interrogatory consists of *at least* three Interrogatories. In
 28 answering the Interrogatory, Meta does not waive this objection.

INTERROGATORY NO. 13:

Describe in detail the policies and procedures that You follow or have followed in assessing whether to permit the use of a dataset as Training Data for a Meta Language Model, and:

- a. list all datasets that have been so reviewed;
- b. list datasets that have been permitted to be used as Training Data for any Meta Language Model, and the date so designated;
- c. list datasets that have at any time been forbidden from being used as Training Data for any Meta Language Model;
- d. for each forbidden dataset, list the reason that it was forbidden and the date so forbidden;
- e. if a forbidden dataset was later designated safe to use, list the reason its designation was revised and the date so revised.

RESPONSE TO INTERROGATORY NO. 13:

Meta incorporates by reference its objections and definitions above, including to the terms “You,” “Training Data,” and “Meta Language Models.”

As an initial matter, Meta objects to this Interrogatory because it consists of at least three Interrogatories, which count toward Plaintiffs’ limit. Specifically, questions about so-called policies or procedures are separate from a list of datasets reviewed, reasons why a dataset was “forbidden” (a term used by the Interrogatory), or reasons why a dataset was once “forbidden” but later not. In answering the Interrogatory, Meta does not waive this objection.

Meta objects to this Interrogatory because, on its face, it does not exclude legal advice or opinions, which are subject to attorney-client privilege and/or attorney work product doctrine, in particular as the subject matter of the Interrogatory encompasses policies and procedures that are legal in nature. Meta will not provide information concerning policies and procedures that constitute legal advice, nor will Meta respond to subparts (a), (b), to the extent that it concerns whether and when a given dataset received approval for use, and (c)-(e).

Meta objects to this Interrogatory as vague and ambiguous as to the term “forbidden dataset,” which is undefined. Meta further objects to this term to the extent it implies that Meta has

1 prohibited use of any particular dataset for training a language model. Meta will construe
2 “forbidden dataset” to mean datasets referred to in subpart (c).

3 Meta objects to this Interrogatory as overbroad, unduly burdensome, and disproportionate
4 to the needs of the case and seeks information that is not relevant to the parties’ claims and defenses,
5 in particular to the extent it seeks information concerning issues unrelated to Plaintiffs’ direct
6 copyright infringement claim.

7 Meta objects to this Interrogatory as duplicative and cumulative of Interrogatory No. 1, and
8 therefore refers to its response to Interrogatory 1.

9 Meta objects to this Interrogatory because it exceeds Plaintiffs’ limit of 25 Interrogatories
10 under Rule 33(a)(1).

11 For these reasons, Meta will not respond to the Interrogatory.

12 **INTERROGATORY NO. 14:**

13 Identify by name all individuals or entities who have applied for access to Llama 1. Include
14 in Your response:

15 a. The names all individuals or entities who obtained access to Llama 1;

16 b. The names all individuals or entities who you denied access to Llama 1, including
17 the names of all individuals or entities who have had their access to Llama 1 revoked or otherwise
18 limited by You;

19 c. For each individual or entity whom you denied access to Llama 1, a description of
20 the reason(s) why you denied access to that particular individual or entity.

21 **RESPONSE TO INTERROGATORY NO. 14:**

22 Meta incorporates by reference its objections and definitions above, including to the terms
23 “You,” “Your,” and “Llama 1.”

24 As an initial matter, Meta objects to this Interrogatory because it consists of at least three
25 Interrogatories, which count toward Plaintiffs’ limit. Specifically, an inquiry regarding entities
26 who applied for access is separate from a question about who obtained or was denied access, and
27 the reasons why (on a person by person basis).

28

Meta objects to this Interrogatory as overbroad, unduly burdensome, and disproportionate to the needs of the case as it would require Meta to compile a list of every individual who sought to use Llama v1 and when and whether they were given access (or not), and then determine the reasoning (if there was one) why access was or was not granted. There is no relevance or proportionality of this request to Plaintiffs' direct copyright infringement claim relating to training.

Meta objects to this Interrogatory because, on its face, it does not exclude legal advice or opinions, which are subject to attorney-client privilege and/or attorney work product doctrine, in particular as the subject matter of the Interrogatory encompasses determinations that are legal in nature.

Finally, Meta objects to this Interrogatory because it exceeds Plaintiffs' limit of 25 Interrogatories under Rule 33(a)(1).

For these reasons, Meta will not respond to the Interrogatory.

INTERROGATORY NO. 15:

Explain in detail the role of the following individuals in any aspect of the design, training, development, testing, marketing, release or support of each of the Meta Language Models:

■	██████████
■	██████████████████
■	██████████████
■	██████████████████
■	██████████████
■	██████████████████
■	██████████
■	██████████████
■	██████████████████
■	██████████
■	██████████████████
■	██████████████

RESPONSE TO INTERROGATORY NO. 15:

Meta incorporates by reference its objections and definitions above, including to the term “Meta Language Models.”

As an initial matter, Meta objects to this Interrogatory because it consists of at least fifteen Interrogatories, which count toward Plaintiffs’ limit. Specifically, each individual’s role in the broad subjects of “design, training, development, testing, marketing, release or support” are separate subjects, and are not necessarily related to one another.

Meta objects to this Interrogatory as vague and ambiguous as to the term “support,” which is undefined and effectively meaningless. Meta also objects to this Interrogatory as vague and ambiguous as to the phrase “involved in any aspect of,” which is undefined and overbroad. Meta will construe this Interrogatory to seek information concerning the identified individuals’ job responsibilities in connection with the “design, training, development, testing, marketing, and release” of the Meta Language Models (as construed above).

Meta objects to this Interrogatory as overbroad, unduly burdensome, and disproportionate to the needs of the case and seeks information that is not relevant to the parties’ claims and defenses, in particular to the extent it seeks information concerning issues unrelated to Plaintiffs’ theory of copyright infringement.

Finally, Meta objects to this Interrogatory because it exceeds Plaintiffs’ limit of 25 Interrogatories under Rule 33(a)(1).

For these reasons, Meta will not respond to the Interrogatory.

Dated: February 23, 2024

COOLEY LLP

By: /s/ Judd Lauter

Judd Lauter
Bobby Ghajar
Mark Weinstein
Kathleen Hartnett
Colette Ghazarian

Attorneys for Defendant
META PLATFORMS, INC.

1 COOLEY LLP
BOBBY GHAJAR (198719)
2 (bghajar@cooley.com)
COLETTE GHAZARIAN (322235)
3 (cghazarian@cooley.com)
1333 2nd Street, Suite 400
4 Santa Monica, California 90401
Telephone: (310) 883-6400
5 MARK WEINSTEIN (193043)
(mweinstein@cooley.com)
6 KATHLEEN HARTNETT (314267)
(khartnett@cooley.com)
7 JUDD LAUTER (290945)
(jlauter@cooley.com)
8 ELIZABETH L. STAMESHKIN (260865)
(lstameshkin@cooley.com)
9 3175 Hanover Street
Palo Alto, CA 94304-1130
10 Telephone: (650) 843-5000

11 LEX LUMINA PLLC
MARK A. LEMLEY (155830)
12 (mlemley@lex-lumina.com)
745 Fifth Avenue, Suite 500
13 New York, NY 10151
Telephone: (646) 898-2055

14 CLEARY GOTTlieb STEEN & HAMILTON LLP
ANGELA L. DUNNING (212047)
15 (adunning@cgsh.com)
1841 Page Mill Road, Suite 250
16 Palo Alto, CA 94304
Telephone: (650) 815-4131

17 *Counsel for Defendant Meta Platforms, Inc.*
18

19 **UNITED STATES DISTRICT COURT**
20 **NORTHERN DISTRICT OF CALIFORNIA**

21 RICHARD KADREY, et al.,
22 Individual and Representative Plaintiffs,
23 v.
24 META PLATFORMS, INC., a Delaware
25 corporation;
26 Defendant.
27
28

Case No. 3:23-cv-03417-VC

**DEFENDANT META PLATFORMS, INC.'S
RESPONSES AND OBJECTIONS TO
PLAINTIFFS' SECOND SET OF
INTERROGATORIES**

Trial Date:
Date Action Filed: July 7, 2023

1 and undue burden and expense. Meta further objects to Instruction 11 to the extent that it purports
 2 to require Meta to investigate information outside of its possession, custody, or control. As such
 3 the Instruction is overly broad, as well. Subject to any objections applicable to a particular
 4 Interrogatory, Meta will conduct a reasonable, proportionate search for non-privileged, relevant,
 5 responsive information within its possession, custody, or control.

6 **19.** In responding to all Interrogatories, Meta will comply with the requirements of the
 7 Federal Rules of Evidence and Federal Rule of Civil Procedure 26.

8 **III. OBJECTIONS AND RESPONSES TO INDIVIDUAL INTERROGATORIES**

9 **INTERROGATORY NO. 1:**

10 State all facts on which you base Your contention that Your conduct constitutes fair use (17
 11 U.S.C. § 107).

12 **RESPONSE TO INTERROGATORY NO. 1:**

13 Meta incorporates by reference its objections and definitions above.

14 Meta objects to this Interrogatory as vague and ambiguous as to the phrase “Your conduct,”
 15 which is undefined and could refer to any conduct. Meta will construe this Interrogatory to seek
 16 information concerning Meta’s claim of fair use in connection with the conduct alleged in the
 17 Complaint (as construed above).

18 Meta objects to this Interrogatory as overbroad, unduly burdensome, and disproportionate
 19 to the needs of the case to the extent it seeks information that Meta does not intend to rely on to
 20 support a claim of fair use and calls for a lengthy narrative with regard to twelve different plaintiffs
 21 and more than forty works.

22 Meta objects to this Interrogatory to the extent it prematurely calls for expert testimony or
 23 identification of facts yet to be disclosed by Plaintiffs, and to the extent that it requires Meta to
 24 respond to legal arguments or theories not yet disclosed by Plaintiffs.

25 Finally, Meta objects to this Interrogatory because it exceeds Plaintiffs’ limit of 25
 26 Interrogatories under Rule 33(a)(1).

27 Subject to and without waiving the foregoing objections, and pursuant to the terms of the
 28 Protective Order, Meta responds as follows:

1 The alleged use of Plaintiffs' asserted works, to the extent it is shown to have occurred, is
 2 highly transformative in nature, because it adds something new, with a further purpose or different
 3 character, altering those works with considerable new expression, meaning, or message. To the
 4 extent Plaintiffs' works were used to train the Meta Language Models, the purpose was
 5 transformative both in terms of purpose and expression. With respect to purpose, Plaintiffs' works
 6 were allegedly used as data to train the models. In other words, the works were allegedly a part of
 7 a corpus of text (specifically, terabytes of text from a variety of sources), from which the models
 8 built complex statistical representations of language derived from the patterns, structures, and
 9 relationships between words within the corpus. This enables the models to predict the next word
 10 in a sequence, and thereby provide useful responses to any manner of input prompts. Such use of
 11 text material serves a fundamentally different purpose from the original texts on which the Meta
 12 Language Models were trained. *See e.g.*, Meta_Kadrey_00000001-00000077,
 13 Meta_Kadrey_00000078-00000104, Meta_Kadrey_00000224-00000248,
 14 Meta_Kadrey_00093669- Meta_Kadrey_00093760.

15 The text corpus used to train the Meta Language Models includes a large amount of textual
 16 materials, and to the extent Plaintiffs' works were used to train those models, they would constitute
 17 a tiny fraction of the textual training dataset (both individually and collectively). The purpose of
 18 the models, and the use of text datasets, is to create new, original textual output, not to reproduce
 19 the content of the datasets with which it was trained. This is demonstrated by, among other things,
 20 Meta's efforts to minimize the models' ability to memorize and/or output training data verbatim
 21 (*see e.g.*, Meta_Kadrey_00000277) and the wide variety of uses that have been made of the models.
 22 *See e.g.*, Meta_Kadrey_00092978-00093308, Meta_Kadrey_00062157. Indeed, all Plaintiffs have
 23 admitted that they are not aware of any output from any Meta Language Model that replicates any
 24 protected expression in their at-issue books. *See e.g.*, T. Coates 7/22/2024 Resps. & Objs. to Meta's
 25 2nd Set of RFAs, Resp. to RFA No. 24 (admitting, subject to objections, that Plaintiff is "personally
 26 unaware of any text generated by any of Meta's Llama models that infringes [Plaintiff's]
 27 ASSERTED WORKS."); J. Díaz 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to
 28 RFA No. 24 (same); C. Golden 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to

1 RFA No. 24 (same); A. Greer 7/22/2024 Resps. & Objs. to Meta’s 2nd Set of RFAs, Resp. to RFA
 2 No. 24 (same); D. Hwang 7/22/2024 Resps. & Objs. to Meta’s 2nd Set of RFAs, Resp. to RFA No.
 3 24 (same); R. Kadrey 7/22/2024 Resps. & Objs. to Meta’s 2nd Set of RFAs, Resp. to RFA No. 24
 4 (same); M. Klam 7/22/2024 Resps. & Objs. to Meta’s 2nd Set of RFAs, Resp. to RFA No. 24
 5 (same); L. Lippman 7/22/2024 Resps. & Objs. to Meta’s 2nd Set of RFAs, Resp. to RFA No. 24
 6 (same); S. Silverman 7/22/2024 Resps. & Objs. to Meta’s 2nd Set of RFAs, Resp. to RFA No. 24
 7 (same); R. Snyder 7/22/2024 Resps. & Objs. to Meta’s 2nd Set of RFAs, Resp. to RFA No. 24
 8 (same); J. Woodson 7/22/2024 Resps. & Objs. to Meta’s 2nd Set of RFAs, Resp. to RFA No. 24
 9 (same); C. Golden 9/17/2024 Dep. 260:6–261:13 (Mr. Golden testifying that the present lawsuit is
 10 not about what comes out of Meta’s large language models); A. Greer 9/24/2024 Rough Drft. Dep.
 11 15:21–16:20, 68:22–69:2 (Mr. Greer testifying that he doesn’t “know anything about” Meta passing
 12 off any of Mr. Greer’s work as AI-generated output, and acknowledging that there are no presently
 13 asserted claims by Plaintiffs as to output generated by Meta’s large language models); D. Hwang
 14 9/16/2024 Dep. 252:23–253:6 (Mr. Hwang testifying that the operative complaint does not allege
 15 that the Meta Language models create any output that is similar to any of his books or plays); L.
 16 TerKeurst 9/23/2024 Dep. 226:20–25, 229:5–12 (Ms. TerKeurst testifying that she was not aware
 17 of any text generated by the Meta Language Models that was substantially similar that of her
 18 asserted works); *cf.* D. Hwang 9/16/2024 Dep. 363:6–15 (Mr. Hwang testifying that he did not
 19 believe when he filed suit that the Meta Language Models could create works that were
 20 substantially similar to his works); L. Lippman 9/17/2024 Rough Drft. Dep. 301:1–7 (Ms. Lippman
 21 testifying that she is aware that the current version of the Complaint does not allege that Meta’s
 22 generative AI tools create any output that is substantially similar to her books).

23 The transformative nature also extends to expression. The pre-training process involves a
 24 number of steps that fundamentally transform the input dataset text in order to facilitate training of
 25 the large language model (“LLM”), which Meta will further describe in further detail in connection
 26 with expert discovery. At a high-level, the pre-training process includes a “tokenization” step in
 27 which the data in training datasets is broken down and encoded into a series of values known as
 28 “tokens” which are used to create numerical vector representations that the LLM training

1 algorithms can understand. (*See e.g.*, Meta_Kadrey_00000078-00000104, Section 2.1
 2 (“Tokenizer”); Meta_Kadrey_00000001-00000077, at 6 (“Tokenizer”).) The input data is then
 3 used in a complex series of LLM training algorithms that adjust the large number of numerical
 4 values (known as parameters which include weights) in the LLM, that define the connections and
 5 relationships between the nodes in the LLM. By adjusting these weights, the LLM can “learn” and
 6 better predict correct outputs based on input data. These numerical parameters enable the LLM,
 7 after the training process, to generate better output data in response to input prompts. The process
 8 of training of an LLM represents a complete transformation into a form that is entirely
 9 unrecognizable from the original training data.

10 As a further indication of the transformative nature, and the fact that the training process
 11 does not simply make a copy of the input dataset, the training process is so computationally
 12 complex that it requires an enormous amount of computing power. *See e.g.*,
 13 Meta_Kadrey_00000001-00000077, at 4 (“When training a 65B-parameter model, our code
 14 processes around 380 tokens/sec/GPU on 2048 A100 GPU with 80GB of RAM. This means that
 15 training over our dataset containing 1.4T tokens takes approximately 21 days.”).

16 Aside from its highly transformative nature, Meta’s alleged use also has substantial non-
 17 commercial, nonprofit, and educational (including research) purposes. Meta is investing [REDACTED]
 18 [REDACTED] in research and development of state-of-the-art LLM technology that it is then making
 19 available to the public. In particular, the training of the Meta Language Models resulted in Meta’s
 20 release of Llama 1, Llama 2, CodeLlama, and Llama 3, 3.1, and 3.2 to the open source
 21 community. These LLMs were provided openly to the public, pursuant to an open license that
 22 permits developers, researchers, and institutions (with the exception of licensees with more than
 23 700 million monthly active users) to use and modify the Llama models free of charge. *See e.g.*,
 24 Meta_Kadrey_00000160-00000162, Meta_Kadrey_00093275-00093284,
 25 Meta_Kadrey_00093658-00093760. The open release of Llama has resulted in the Llama models
 26 being downloaded hundreds of millions of times by researchers and developers from around the
 27 world and has catalyzed development of new and improved AI tools and technologies. *See e.g.*,
 28 Meta_Kadrey_00092978-00093308, Meta_Kadrey_00062157. The open and public release of

1 Llama has allowed the public to access highly capable LLM technologies that would otherwise be
2 available only to large organizations and/or at considerable expense. More broadly, Meta's
3 investment and open release is contributing to the U.S. economy, the emergence of a new and
4 important industry, and the U.S.'s global leadership of that industry over geopolitical competitors.

5 Plaintiffs allegedly infringed works were published prior to Meta's alleged use of those
6 works. And Meta's use of large volumes of texts was necessary to extract data regarding, e.g.,
7 word frequencies, grammar, and syntax from those works to generate new content. This
8 information constitutes either facts and ideas, which are not protectable by copyright, or is
9 otherwise unrelated to the purpose of copyright protection. Furthermore, given that training the
10 Meta Language Models requires terabytes of text, that greater volumes of text tends to improve
11 model performance on objective benchmarks measuring reasoning and knowledge of facts, and the
12 formats in which the text is available, it was reasonable for Meta to utilize copies of entire works
13 (as opposed to portions thereof), in particular given Meta's efforts to develop the models in a
14 manner that minimizes the likelihood that training data can be reproduced as model output. For
15 example, books data comprised only approximately 4.5% of tokens used to train Llama 1 and [REDACTED]
16 [REDACTED] of tokens used to train Llama 2.

17 The ordinary market for Plaintiffs' works is the market for people to purchase and read the
18 books and, possibly, to create derivative works from those books, such as for audio books and film
19 and television adaptations. The Meta Language Models and their outputs do not serve as a market
20 substitute for the Plaintiffs' asserted works, do not compete with those works, and do not harm the
21 value of Plaintiffs' asserted works. Plaintiffs have produced no evidence to the contrary, such as
22 evidence of lost sales or other financial harm. To the contrary, plaintiffs have admitted that they
23 are not aware of any such harm. *See e.g.*, T. Coates 9/6/2024 Resps. & Objs. to Meta's 2nd Set of
24 RFAs, Resp. to RFA No. 15 (admitting, subject to objections, that Plaintiff is unaware of lost sales
25 due to alleged infringement); J. Diaz 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to
26 RFA No. 15 (same); C. Golden 8/28/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to
27 RFA No. 15 (same); A. Greer 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA
28 No. 15 (same); D. Hwang 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No.

15 (same); R. Kadrey 8/28/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 15 (same); M. Klam 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 15 (same); L. Lippman 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 15 (same); S. Silverman 8/28/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 15 (same); R. Snyder 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 15 (same); J. Woodson 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 15 (same); L. Terkeurst 9/12/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 15 (similar); A. Greer 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 16 (admitting, subject to objections, that Plaintiff's book sales have not decreased due to the alleged use of Plaintiff's Asserted Works to train large language models); L. Terkeurst 9/6/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 16 (similar); T. Coates 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (admitting, subject to objections, that Plaintiff is unaware of lost licensing opportunities due to alleged infringement); J. Diaz 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); C. Golden 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); A. Greer 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); D. Hwang 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); R. Kadrey 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); M. Klam 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); L. Lippman 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); S. Silverman 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); R. Snyder 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); J. Woodson 9/19/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); L. Terkeurst 9/12/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA No. 18 (same); T. Coates 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (admitting, subject to objections, that Plaintiff is unaware of (1) persons reading text generated by Meta's Llama models as a substitute for Plaintiff's Asserted Works as described in RFA 22, and (2) documentary evidence of persons reading text generated by Meta's Llama models as substitute for Plaintiff's Asserted Works as described in RFA 23); J. Diaz

7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); C. Golden 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); A. Greer 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); D. Hwang 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); R. Kadrey 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); M. Klam 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); L. Lippman 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); S. Silverman 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); R. Snyder 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); J. Woodson 7/22/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (same); L. TerKeurst 8/21/2024 Resps. & Objs. to Meta's 2nd Set of RFAs, Resp. to RFA Nos. 22 and 23 (similar); D. Hwang 9/16/2024 Dep. 254:7–13 (Mr. Hwang testifying that he did not know whether he had experienced any financial loss as a result of the alleged training of the Meta Language Models on his asserted works); D. Hwang 9/16/2024 Dep. 290:10–17, 291:22–292:4 (Mr. Hwang testifying that he was not aware of any lost sales of his asserted works due to the alleged infringement in the complaint); R. Kadrey 9/25/2024 Rough Drft. Dep. 194:10–21 (Mr. Kadrey testifying that he was unaware of any injury other than purported harm of not receiving compensation from Meta); R. Kadrey 9/25/2024 Rough Drft. Dep. 193:4–7 (Mr. Kadrey testifying that he was not aware of any lost sales of his asserted works due to the alleged infringement in the complaint); L. Lippman 9/17/2024 Rough Drft. Dep. 329:19–330:5 (Ms. Lippman testifying that she is unaware of any financial harm that she has suffered as a result of conduct by Meta alleged in the Complaint); A. Greer 9/24/2024 Rough Drft. Depo. 110:3–20 (Mr. Greer testifying that he is unaware of any lost book sales or lost licensing opportunities as a result of Meta's alleged infringement of his works, or whether it is possible that book sales have actually increased due to his participation in this lawsuit).

Moreover, there was no market for licensing Plaintiffs' literary books as training data for LLMs at the time Meta first Llama 1 or at the time Plaintiffs' filed their lawsuit, and there is no such market today. Plaintiffs have admitted as much in discovery. Moreover, any particular book

1 has no independent value as training data and/or is interchangeable with countless other books. In
2 any case, such a market would be for a transformative use. It would also be impractical, if not
3 impossible, for companies developing LLMs to attempt to negotiate licenses with each individual
4 book rights holder for various reasons, including but not limited to the amount of time and cost
5 necessary to do so would have precluded development of the models in the first instance, in
6 particular given the time and costs relative to the de minimis value of individual works to the
7 development of the models. In support, Meta intends to rely on Plaintiffs' discovery responses and
8 testimony, expert reports and testimony, percipient witness testimony, as well as documents
9 produced in this litigation.

10 **INTERROGATORY NO. 2:**

11 If You or any of Your employees and/or agents intend to assert the advice of counsel
12 defense, state any and all facts upon which You or any of your employees and/or agents intend to
13 rely on for that contention.

14 **RESPONSE TO INTERROGATORY NO. 2:**

15 Meta incorporates by reference its objections and definitions above.

16 Meta objects to this Interrogatory as vague and ambiguous as to the reference to "Your
17 employees and/or agents" with respect to any defense in this case, as no Meta employees or agents
18 are parties to this case.

19 Meta objects to this Interrogatory because it exceeds Plaintiffs' limit of 25 Interrogatories
20 under Rule 33(a)(1).

21 Subject to and without waiving the foregoing objections, Meta responds as follows: Meta
22 does not presently intend to assert the advice of counsel defense.

1 Dated: September 30, 2024

COOLEY LLP

2 By: /s/ Judd Lauter

3 Bobby Ghajar
4 Mark Weinstein
5 Kathleen Hartnett
6 Judd Lauter
7 Liz Stameshkin
8 Colette Ghazarian

9 LEX LUMINA PLLC
10 Mark A. Lemley

11 CLEARY GOTTlieb STEEN &
12 HAMILTON LLP
13 Angela L. Dunning

14 Attorneys for Defendant
15 META PLATFORMS, INC.

BOIES SCHILLER FLEXNER LLP

David Boies (*pro hac vice*)
333 Main Street
Armonk, NY 10504
(914) 749-8200
dboies@bsflp.com

Maxwell V. Pritt (SBN 253155)
Joshua I. Schiller (SBN 330653)
Joshua M. Stein (SBN 298856)
44 Montgomery Street, 41st Floor
San Francisco, CA 94104
(415) 293-6800
mpritt@bsflp.com
jischiller@bsflp.com
jstein@bsflp.com

Jesse Panuccio (*pro hac vice*)
1401 New York Ave, NW
Washington, DC 20005
(202) 237-2727
jpanuccio@bsflp.com

David L. Simons (*pro hac vice*)
55 Hudson Yards, 20th Floor
New York, NY 10001
(914) 749-8200
dsimons@bsflp.com

JOSEPH SAVERI LAW FIRM, LLP

Joseph R. Saveri (SBN 130064)
Cadio Zirpoli (SBN 179108)
Christopher K.L. Young (SBN 318371)
Holden Benon (SBN 325847)
Aaron Cera (SBN 351163)
601 California Street, Suite 1505
San Francisco, California 94108
(415) 500-6800
jsaveri@saverilawfirm.com
czirpoli@saverilawfirm.com
cyoung@saverilawfirm.com
hbenon@saverilawfirm.com
acera@saverilawfirm.com

Matthew Butterick (SBN 250953)
1920 Hillhurst Avenue, #406
Los Angeles, CA 90027
(323) 968-2632
mb@buttericklaw.com

CAFFERTY CLOBES MERIWETHER & SPRENGEL LLP

Bryan L. Clobes (*pro hac vice*)
135 S. LaSalle Street, Suite 3210
Chicago, IL 60603
(312)-782-4880
bclobes@caffertyclobes.com

*Counsel for Individual and Representative Plaintiffs
and the Proposed Class. (additional counsel included below)*

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION**

Richard Kadrey, et al.,

Individual and Representative Plaintiffs,

v.

Meta Platforms, Inc.,

Defendant.

Lead Case No. 3:23-cv-03417-VC

**PLAINTIFFS' THIRD SET OF
INTERROGATORIES TO DEFENDANT
META**

1 response is or has become in some material respect incomplete or incorrect, pursuant to Federal Rule of
2 Civil Procedure 26(e).

3 10. To be clear, each interrogatory seeks information from all predecessors, successors,
4 subsidiaries, departments, divisions, and/or affiliates of Meta.

5 **INTERROGATORIES**

6 **INTERROGATORY NO. 19:**

7 Describe the total revenue You have received and forecast or otherwise expect you will receive
8 from the distribution or use of Llama Models since the start of their development, including by way of
9 any direct sales, licensing fees, or other financial or in-kind benefits.

10 **INTERROGATORY NO. 20:**

11 Describe any partnerships or collaborations related to the Llama Models You have entered with
12 third parties that have resulted in financial benefits, including the details of such arrangements, such as
13 financial terms, revenue or in-kind benefits received in relation to them.

14 **INTERROGATORY NO. 21:**

15 Identify all Person(s) involved in the decision to use Shadow Datasets to train Llama Models.

16 **INTERROGATORY NO. 22:**

17 Describe any efforts You have made to obtain licenses or any similar permissions to use
18 Shadow Datasets, or the works contained therein, to train Llama Models.

19 **INTERROGATORY NO. 23:**

20 Identify all sources from which You have obtained Shadow Datasets.

21 **INTERROGATORY NO. 24:**

22 Identify all individuals, including current and former employees, who work or have worked on
23 securing licenses for data or material You have used, have planned to use, or plan to use to train Llama
24 Models.

25 **INTERROGATORY NO. 25:**

26 Identify all Meta employees who communicated with third parties—including Microsoft
27 Corporation, OpenAI, Eleuther AI, Anthropic, Dell Inc., Cloudflare, Inc. and Qualcomm—regarding
28 data used to train Llama Models.

INTERROGATORY NO. 26:

Identify all individuals involved in the process of searching for, identifying, collecting, and producing responsive documents and data in this matter, and their roles in that process.

INTERROGATORY NO. 27:

Describe all revenue sources related to the Llama Models, specifying and quantifying the amounts received from each source, including but not limited to sales, licensing, partnerships, and any other financial or in-kind arrangements.

INTERROGATORY NO. 28:

Describe the profits associated with Your distribution, use or provision of the Llama Models, including a detailed explanation of the calculation of the profits.

INTERROGATORY NO. 29:

Describe any internal analyses, reports, or forecasts You conducted or prepared regarding the financial impact of the Llama Models on Your revenue and profitability since development of the Llama Models, including analyses, reports, and forecasts concerning future revenue and profitability.

INTERROGATORY NO. 30:

Identify and describe all costs You incurred in the development, training, and distribution of the Llama Models, including but not limited to data acquisition, use of computational resources, personnel costs, and marketing costs.

INTERROGATORY NO. 31:

Identify and describe any internal analyses, reports, forecasts, or estimates of the costs associated with licensing data for use in the development and training of the Llama Models.

INTERROGATORY NO. 32:

Describe how Third-Party applications or services have generated revenue for You through the use of the Llama Models, including the financial or in-kind benefits received from each use.

INTERROGATORY NO. 33:

Describe Your receipt of research grants or funding related to the development, use or provision of the Llama Models, including the amounts received and the sources of such funding.

INTERROGATORY NO. 34:

1 Describe any Communications between Meta and EleutherAI regarding the use of Shadow
2 Datasets.

3 **INTERROGATORY NO. 35:**

4 Describe any legal review or analysis regarding the use of the Shadow Datasets to train Llama
5 Models.

6 **INTERROGATORY NO. 36:**

7 Identify all individuals or entities who had access to the Llama Models' source code after their
8 release.

9 **INTERROGATORY NO. 37:**

10 Identify all guidelines and policies established to prevent disclosure of confidential
11 communications or proprietary data concerning Llama Models.

12 **INTERROGATORY NO. 38:**

13 Identify the Meta-controlled physical location(s) where Plaintiffs may access current Llama
14 Model versions for inspection and analysis.

15 **INTERROGATORY NO. 39:**

16 Describe the criteria used to determine who would receive access to the Llama Models.

17 **INTERROGATORY NO. 40:**

18 Describe any measures taken to ensure databases You chose for training Llama Models did not
19 include copyrighted material.

20 **INTERROGATORY NO. 41:**

21 Describe any measures taken to remove or filter copyrighted material out of databases You
22 chose for training Llama Models.

23 **INTERROGATORY NO. 42:**

24 Describe any complaints or legal notices You have received regarding the use of copyrighted
25 material to train the Llama Models.

26 **INTERROGATORY NO. 43:**

27 Describe any analysis or assessment undertaken to identify or quantify potential legal risks
28 associated with using the Shadow Datasets.

INTERROGATORY NO. 44:

Describe the purposes for which You use SRT, including whether it is used for business operations, compliance monitoring, or any other non-legal purposes or activities.

INTERROGATORY NO. 45:

Identify all individuals or departments with access to SRT and describe the ways they use it, including whether it is used for business operations, compliance monitoring, or any other non-legal purposes or activities.

INTERROGATORY NO. 46:

Identify any communications, messages, or data related to Llama Models that were entered into or tracked by SRT.

INTERROGATORY NO. 47:

Describe the methods used to identify and collect relevant documents and information in this matter, including custodial files searched, data or document repositories searched, and search terms employed and hit reports generated.

INTERROGATORY NO. 48:

Describe the contents of the data sets Meta has produced in this matter, including those labeled Meta_Kadrey_Data_001, Meta_Kadrey_Data_002, Meta_Kadrey_Data_003, Meta_Kadrey_Data_004, and Meta_Kadrey_Data_005.

Dated: October 10, 2024

By: /s/ David L. Simons

JOSEPH SAVERI LAW FIRM, LLP

Joseph R. Saveri (SBN 130064)
 Cadio Zirpoli (SBN 179108)
 Christopher K.L. Young (SBN 318371)
 Holden Benon (SBN 325847)
 Aaron Cera (SBN 351163)
 601 California Street, Suite 1505
 San Francisco, California 94108
 (415) 500-6800
 jsaveri@saverilawfirm.com

BOIES SCHILLER FLEXNER LLP

David Boies (*pro hac vice*)
 333 Main Street
 Armonk, NY 10504
 (914) 749-8200
 dboies@bsflp.com

Maxwell V. Pritt (SBN 253155)
 Joshua I. Schiller (SBN 330653)
 Joshua M. Stein (SBN 298856)

1 czirpoli@saverilawfirm.com
 2 cyoung@saverilawfirm.com
 3 hbenon@saverilawfirm.com
 4 acera@saverilawfirm.com

44 Montgomery Street, 41st Floor
 San Francisco, CA 94104
 (415) 293-6800
 mpritt@bsflp.com
 jischiller@bsflp.com
 jstein@bsflp.com

5 Matthew Butterick (SBN 250953)
 6 1920 Hillhurst Avenue, #406
 7 Los Angeles, CA 90027
 8 (323) 968-2632
 9 mb@buttericklaw.com

Jesse Panuccio (*pro hac vice*)
 1401 New York Ave, NW
 Washington, DC 20005
 (202) 237-2727
 jpanuccio@bsflp.com

**CAFFERTY CLOBES MERIWETHER &
 SPRENGEL LLP**

Bryan L. Clobes (*pro hac vice*)
 135 S. LaSalle Street, Suite 3210
 Chicago, IL 60603
 (312)-782-4880
 bclobes@caffertyclobes.com

David L. Simons (*pro hac vice*)
 55 Hudson Yards, 20th Floor
 New York, NY 10001
 (914) 749-8200
 dsimons@bsflp.com

DICELLO LEVITT LLP

Amy Keller (*pro hac vice*)
 Nada Djordjevic (*pro hac vice*)
 James Ulwick (*pro hac vice*)
 (312) 214-7900
 akeller@dicellolevitt.com
 ndjordjevic@dicellolevitt.com
 julwick@dicellolevitt.com
 10 North Dearborn Street, Sixth Floor
 Chicago, Illinois 60602

David Straite (*pro hac vice*)
 (646) 933-1000
 dsraite@dicellolevitt.com
 485 Lexington Avenue
 Suite 1001
 New York, New York 10017

CERTIFICATE OF SERVICE

I, the undersigned, am employed by Boies Schiller Flexner LLP. My business address is 55 Hudson Yards, New York, NY 10001. I am over the age of eighteen and not a party to this action.

On October 10, 2024, I caused the following documents to be served by email upon the parties listed on the attached Service List:

- **PLAINTIFFS' THIRD SET OF INTERROGATORIES TO DEFENDANT META PLATFORMS, INC.**

I declare under penalty of perjury that the foregoing is true and correct. Executed October 10, 2024, at New York, New York.

By: /s/ David L. Simons

SERVICE LIST

Bobby A. Ghajar
Colette Ani Ghazarian
COOLEY LLP
1333 2nd Street, Suite 400
Santa Monica, CA 90401
Email: bghajar@cooley.com
cghazarian@cooley.com

Kathleen R. Hartnett
COOLEY LLP
3 Embarcadero Center, 20th Floor
San Francisco, CA 94111-4004
Email: khartnett@cooley.com

Judd D. Lauter
Elizabeth Lee Stameshkin
COOLEY LLP
3175 Hanover Street
Palo Alto, CA 94304
Email: jlauter@cooley.com
lstameshkin@cooley.com

Mark Alan Lemley
LEX LUMINA PLLC
745 Fifth Avenue, Suite 500
New York, NY 10151
Email: mlemley@lex-lumina.com

Angela L. Dunning
**CLEARY GOTTlieb STEEN &
HAMILTON LLP**
1841 Page Mill Road
Palo Alto, CA 94304-1254
Email: adunning@cgsh.com

*Counsel for Defendant
Meta Platforms, Inc.*